Title: Automated insect monitoring with camera traps is transforming ecological understanding

Authors: Mark A.K. Gillespie[1,2,*], Kim Bjerge[3], Jamie Alison[1], Hannah Sørine Gerlich[1], Flemming Helsing-Nielsen[4], Guillaume Mougeot[1], Asger Svenning[1], Simon F.A. Wogram[6,7], Toke T. Høye[1,5]

Affiliations:

[1]Department of Ecoscience, Aarhus University, C.F. Møllers Allé 8, 8000 Aarhus C, Denmark

[2]Department of Civil Engineering and Environmental Science, Western Norway University of Applied Sciences, Sogndal, Norway

[3]Department of Electrical and Computer Engineering, Aarhus University, Finlandsgade 22, 8200 Aarhus N, Denmark

[4]NaturConsult, Skrænten 5, 9520 Skørping, Denmark

[5]Arctic Research Centre, Aarhus University, 8000 Aarhus C, Denmark

[6]Department of Community Ecology, UFZ-Helmholtz Centre for Environmental Research, Halle, Germany

[7]Institute of Biology, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

*Corresponding author: Mark Gillespie, email: markg@ecos.au.dk

Abstract

Addressing global declines in insect biodiversity requires both ecological restoration and high-quality monitoring data. While long-term participatory schemes have been foundational, recent advances in automated recording and AI-based identification offer transformative but undocumented potential. Here, we show how leveraging insect camera traps, deep learning models and statistics drives a step-change in ecological knowledge. We highlight four key areas of ecological understanding: phenology, abundance, richness, and community dynamics, and show how automated data can correct phenological estimates by weeks and improve biodiversity assessments. Data from insect camera traps offer unprecedented resolution and scalability, making them powerful tools for tracking insect communities and informing conservation strategies.

1. Introduction

The recent well documented changes in global biodiversity have drawn concern across the conservation community [1], prompting a concerted effort to improve data collection and deepen understanding of these rapid transformations [2–5]. High impact reports of declines in insect abundance [6,7], biomass [8] and diversity [9], likely driven by myriad anthropogenic pressures [1], are alarming not least due to their impact on ecosystem functioning and human well-being [10,11]. Efforts to reduce or even reverse these trends are anchored by agreed international objectives, such as the Kunming-Montreal Global Biodiversity Framework [12]. However, assessing the progress of the international community in meeting these targets is only possible with high-quality data [13,14], of which there is a lamentable paucity for many insect groups [15,16]. Such long-term, standardised and high-resolution data are essential to advancing ecological research and informing conservation priorities [2,14,17,18].

Collecting such gold-standard data on insect populations and communities has traditionally been challenging. Long-term datasets on insect abundance exist in a few countries such as the UK and the Netherlands, and have provided important insights into declines in butterflies [19,20] and moths [16,21], as well as into the wide range of species-specific responses to anthropogenic drivers and climate variability. To some extent, a broader perspective on insect biodiversity trends can also be achieved by combining long-term datasets covering a variety of taxa and regions [22]. However, these monitoring initiatives have been developed incrementally over decades, are supported by a dedicated network of volunteers in the absence of consistent funding, and boast many other features of an effective monitoring system [5,14]. Establishing comparable data streams in new regions will require great effort and the rapid development of scalable and standardised monitoring protocols that can be implemented globally [23,24]. Such protocols are key to providing comparable biodiversity indicators, such as the Essential Biodiversity Variables (EBVs) suggested by the Group on Earth Observations Biodiversity Observation Network (GEO BON; [25,26].

Automated monitoring systems can make a major contribution to meeting these challenges in the near future, and have been developing rapidly [24,27,28]. For example, advances in DNA metabarcoding have enabled rapid taxonomic identification from the preservative that samples of specimens are stored in [29], bulk-trapped specimens [30], or from water and soil samples [31,32]. Similarly, the use of automatic sensors such as those capturing visual or audio data enables remote collection of data across multiple locations simultaneously [33,34]. Such technological innovations have enormous potential to increase the quantity and quality of data available to ecologists and conservationists, both spatially and temporally [2,27]. For the taxonomic groups on which they are trained, automated systems can deliver information on key EBVs such as the abundance and distribution of individual species, community diversity and composition and the timing of key life history events at high temporal resolution [24,27,35]. Furthermore, emerging evidence suggests that data generated through automated monitoring may provide species richness, abundance and trend information comparable to traditional destructive sampling methods [36,37].

The promise of such data is both exciting and daunting. Insect species richness is many orders of magnitude greater than that of mammals and birds, and taxonomic identification of automatically collected data (i.e., photographs, videos, or sounds recordings) is typically limited by the availability of skilled experts. Furthermore, advances in autonomous operation of the monitoring system, such that species identification is incorporated in the process, have suffered from three main constraints:

1) **Hardware development** has often focused on ad hoc combinations of microcomputers and cameras with too little attention to data management, reproducibility and durability.
2) **Deployment efforts** have generally lacked the scale and ambition needed to enable meaningful comparisons across regions and seasons.
3) **Image analysis workflows and machine learning classifiers** have not yet achieved the robustness required to allow for efficient, rigorous and comprehensive processing of images collected from insect camera traps [38].

However, one approach to the automated monitoring of nocturnal insects has begun to lift these constraints [34,39–41]. The development and widespread deployment of the Automated Monitoring of Insects (AMI) trap system have now generated sufficient data to capture ecological signals of relevance to conservation at scale. As a result, we are now able to begin visualising the transformative potential of this technology. In this paper, we highlight some of the key insights that an automated insect monitoring system can provide to ecologists, decision makers and society in general by sharing initial results of three years of data collection with twelve AMI traps across three regions of Denmark.

2. Automated monitoring of insects (AMI) traps

The AMI trap was developed to monitor nocturnal insects, but the automatic species identification aspect of the system is particularly well developed for moths. Night-active moths have been well studied because of their ease of capture, diverse life histories and the tight links between community composition and variation in local habitat and plant diversity [42,43]. Similarly, moth biodiversity tends to closely reflect even fine-scale changes in environmental factors and management gradients, promoting the group as candidate bioindicator taxa [42–44]. As their responses to drivers of change may also reflect changes in other species groups and ecosystem functioning [16], moths are an ideal target group for a scalable and standardised data collection system.

Some of the traditional moth biodiversity monitoring schemes and studies have already given us high quality information for certain taxa and regions. For example, the Rothamsted Insect Survey has tracked insect distributions across the UK since the 1960s, revealing declines in approximately two-thirds of the studied macro-moth species [45]. Similarly, organised data collection in the Netherlands provided similar findings between 1980 and 2009 [46], while data from the National Moth Monitoring scheme in Finland revealed spatially variable patterns in abundance, richness and population trends [47,48]. However, most assessments are based on data from voluntary monitoring schemes, which provide patchy data often with low temporal resolution, limited standardization and bias towards geographically accessible areas [5]. These limitations are inevitable in voluntary schemes, because participants cannot be expected to monitor plots daily or to brave all kinds of weather. While data imbalances can to some extent be dealt with using appropriate modelling techniques [49], we demonstrate below that inference from traditional insect monitoring may be biased due to the limited temporal resolution and the requirement to sample only during favourable weather conditions.

The design and functionality of the AMI moth traps have been detailed elsewhere [39,40], but briefly, a UV light attracts adult night-flying moths to the trap location, and they land on a white screen towards which a web camera is focussed. The camera is assisted by a light source that ensures diffuse illumination and is connected to a Raspberry Pi 4 micro-computer installed with a motion program that captures images whenever the movement of insects above a predetermined size (~40 mm^2) is detected on the white screen between 11pm and 3am. The motion-triggered images, saved every 2 seconds, are supplemented with time-lapse images saved every 10 minutes. Together, this often results in >6 000 images per night, which are stored on a portable SSD hard drive. The system is powered by a 12V battery and can be accompanied by a solar panel to ensure continuous recording through the summer [40].

The images from the cameras are processed by a deep learning algorithm (Moth Classification and Counting; Open source code: https://github.com/kimbjerge/MCC24-trap), which automatically detects, identifies and counts known moth species using a classifier trained on GBIF data of 2530 moth species found in the UK and Denmark [41,50]. As an individual moth can remain on the white sheet across several images, "tracking" is also used to follow the movement of individuals and reduce multiple observations of the same individual [39]. While a moth could leave and return to the trap later to be counted as a new individual, tracking is a step toward prevention of double-counting. Nevertheless, what the system captures is best understood as an index of relative rather than absolute abundance. Although activity is often closely related to abundance, the strength of this relationship can vary among taxa and contexts, and comparative studies with established approaches such as capture–mark–recapture are still lacking. We therefore use the term "relative abundance" in the sense of a standardized activity index that allows comparisons across space and time, while acknowledging its current limitations.

The dataset analysed here includes data from 12 AMI traps, four in each of three managed nature reserve sites in Denmark: Lille Vildmose (hereafter "LV"), Ovstrup Hede ("OH"), and Søholt Storskov ("SS"). Traps in a site were spaced between 1 and 10 km apart in a variety of habitats including bogs, heathlands and forests. The traps recorded images from April to October in 2022, 2023 and 2024. In 2022 alone, the traps were in operation for 1399 nights, capturing over 3.4 million images containing over 45.2 million instances of insects, comprising over 1.1 million valid insect "tracks". The moth species classifier distinguished over 1600 species, although model performance was evaluated using 47 905 images of only 497 species that were verified by a Danish moth taxonomist. We found average classification precision values of 66.4% (macro - average of precision across all classes) and 77.0% (micro - total number of true positives divided by the total number of classifications). However, to reduce computation time for some of the analyses in this paper, we focused on the observations of a target list of species that were classified with the best precision (detailed below and in Methods).

3. Ecological knowledge from high quality data

In their review of biodiversity monitoring best practice, Montgomery et al. [5] highlight that monitoring data should at least capture one of four key attributes of biodiversity: 1) occurrence and distribution, 2) phenology, 3) abundance and 4) species diversity and community composition. These are also among the suggested EBVs by GEO BON for future global biodiversity monitoring programmes [25,26]. In this section, we demonstrate how the AMI traps provide data for three of these four attributes, outlining analyses for community and species phenology (section 3.1), community and species abundance (3.2), species richness (3.3) and community composition (3.4). In parallel, we highlight the key advantages of the traps over traditional methods of data collection.

3.1 Community- and species-level phenology

For many species of insect, plant, bird and mammal, phenology is advancing with climate change because the climatic cues that trigger life history events (e.g., emergence, bud burst, migration) are occurring earlier [51]. These changes can have species level impacts, but can also affect the interactions between species, particularly as there is much variability in the degree to which species respond to these changes [52]. Studies of moths have shown considerable variation in phenological response to climate change, sometimes linked to species traits such as larval diet, seasonality or overwintering stage [53–56]. However, these findings come from a range of data collection methods and there is a pressing need for more standardised, high-frequency data to fully capture the complexity and variability of insect phenology [5].

The AMI traps are uniquely positioned to meet this need, offering dense intra-annual replication and extremely consistent data collection across years to detail community- and species-level alterations to flight seasons. To demonstrate this, we used Generalised Additive Modelling (GAMs) to model the daily abundance of the 100 most common and most precisely identified species in the dataset, both as a community and individually (see Box 1). This technique fits a smoothed curve to the abundance data over time to represent the phenological pattern of the community and individual species. The phenology curves can then be further analysed to identify key phenological events such as the onset, peak and end of adult activity. Using the three years of data, the community peak phenology and length and shape of the flight season curve exhibited high interannual variability (Fig. 1a). For example, the bimodal seasonal cycle reported for other moth communities [42], is only present in two out of three years, and the timing of the spring peak varies considerably. Similarly, species-level phenological events demonstrate substantial interspecific variability across years, but are relatively consistent within species from year to year (Fig. 1b). Species level differences in phenology and phenological shifts are likely to be a function of varied responses to weather variables, but may also reflect species interactions, dynamic inter-annual effects or trait-based responses to changing site level factors [16,45,56]. Future statistical analysis of this type of data collected over longer time periods will help to pinpoint the drivers behind the sensitivity of species to a range of drivers.

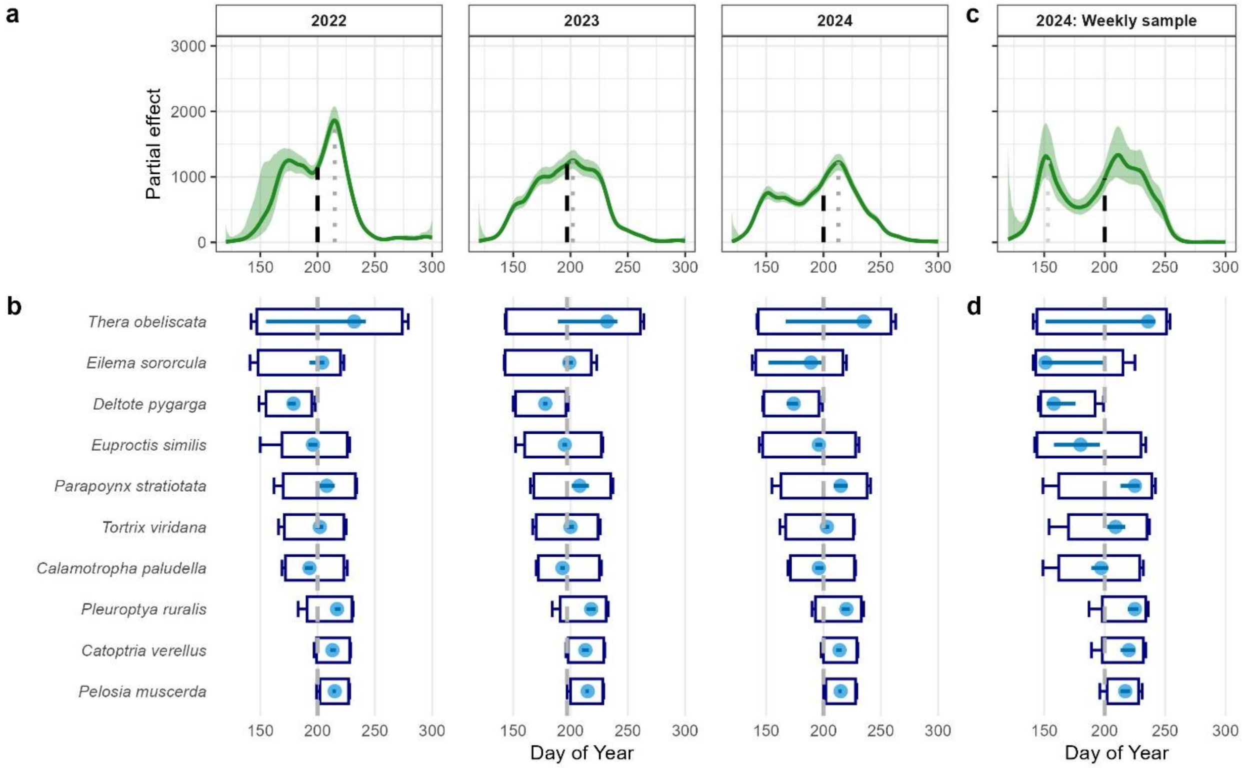


Figure 1: Moth phenology metrics from the AMI traps estimated by Generalised Additive Modelling: a) the estimated community-level distribution curves based on the most abundant and precisely identified 100 species for the three years, with the peak phenology (50% cumulative abundance) highlighted by a dashed line, and maximum abundance by a dotted line; b) the phenology of the 10 most abundant species estimated: left edge of box = activity onset date, right edge of box = activity end date, central dot = peak phenology date. Uncertainty in these three dates is shown by the box whiskers and the central line overlapping the dots; c) For 2024 only, we drew a “weekly sample” using the data from every seventh day and repeated the model to demonstrate the increase in uncertainty and shift in peak phenology in the community curve, and d) in the species level metrics.

To quantify the added value of daily abundance data, we constructed and compared similar estimates with a weekly subsample of data. This would replicate the more typical sampling interval of traditional monitoring approaches, by using only data from every seventh day to quantify changes in phenology estimates and their uncertainty. This subsampled dataset provided estimates with considerably greater uncertainty and different peak phenology dates (shown in Fig 1c & d for 2024 only). Furthermore, at the community level, the bimodal nature of the community curve was exaggerated, while uncertainty at the beginning of the season was particularly high (Fig. 1c). At the species level, some estimates of peak phenology differed markedly, certain activity seasons appeared to become elongated, and uncertainty often increased (Fig. 1d). For many species, the additional information provided by daily observations is critical to producing precise estimates of vital phenological metrics over time, and accurately track changes in phenology [5]. This will be pertinent for rarer species or those with short flight periods, which may not be captured frequently enough to determine phenological patterns using traditional methods, particularly if several consecutive nights of optimal weather conditions are needed to detect activity. Furthermore, combining daily, and potentially even hourly data with weather variables collected at each trap location will provide highly detailed estimates of species and community responses to climate. Such data are highly sought after as the basis for more realistic predictions from mechanistic models [57].

3.2 Relative abundance

In addition to tracking changes in phenology, recording total and species-level abundance is a desirable objective of monitoring systems [5]. Abundance is arguably the most recognisable metric for non-ecologists and is typically linked with ecosystem function and with the success or failure of conservation and management actions [1], or with changes in environmental conditions [8]. For example, data from the UK National Moth Recording Scheme revealed a 31% decline in total moth abundance over 35 years from 1968, with significant declines particularly in the southern half of the country where habitat degradation has been more pronounced [45]. However, spatial and species-specific trends in these declines highlight the importance of local environmental factors and species-level sensitivities to such environmental factors [16,45–47].

While no survey method can produce perfect measures of abundance, total and species-level relative abundance trends can be approximated using the phenology curves shown in Fig 1. The area under the curve offers a practical abundance index that accounts for seasonal variation, sampling effort and the amount of time the species is active [58] (see Box 1). These indices provide standardised inter-annual comparisons and may help mitigate potential biases from double counting in automated tracking systems. Such estimates are widely used for butterfly monitoring data [58], and offer considerable improvements over single-date snapshots of abundance which are highly sensitive to sample timing [59]. However, it should be reiterated that such indices are only comparable across similar trapping regimes, due to limited understanding of the relationship between activity and abundance.

To further illustrate the potential of automated trapping, we computed relative abundance indices for our dataset of 100 species recorded daily, as well as the weekly subsampled dataset described above. For this demonstration we also used a third “good weather” dataset, for which we extracted only the observations collected on nights with temperatures above 12°C, wind speed below 1 m/s, and no rain. This latter dataset was intended to represent a more traditional survey regime, where researchers or volunteers must ensure maximum insect activity before deciding to sample. In addition, we explored the impact of incorporating trap-level weather conditions (temperature, wind speed, relative humidity and a covariate to account for general moth activity *sensu* [60]) into our models to estimate abundance while accounting for weather effects on activity. Thus, six GAMs were constructed in total: models with and without weather variables for the full daily dataset, the weekly dataset and the good weather dataset.

At the community level, abundance index estimates from the daily dataset were generally lower than the Weekly and Good weather samples (Fig. 2a), suggesting that more coarse resolution datasets may overestimate abundance when estimated this way. This difference in estimates may occur because rarer species and those with short flight periods are more often missed during Weekly or Good weather sampling regimes, and the higher relative abundance estimates represent a bias towards common, abundant species. Furthermore, there were differences in trend and uncertainty among the data types, and models that excluded weather variables tended to produce higher estimates of total abundance in most years. At the species level, we focussed on the differences in estimates between models with and without weather variables, and a selection of contrasting patterns demonstrate the potential impact of including or excluding weather conditions from abundance estimates (Fig. 2b). Moth species are well known to respond variously to different weather variables [61]; for some apparently weather sensitive species, models without weather variables produced higher abundance estimates, greater uncertainty and sometimes opposite trends (e.g. *Deltote pygarga*, Marbled White Spot). Conversely, less sensitive species tended to exhibit index estimates that were unaffected by the inclusion of weather variables (e.g. *Anania lancealis*, the Long Winged Pearl).

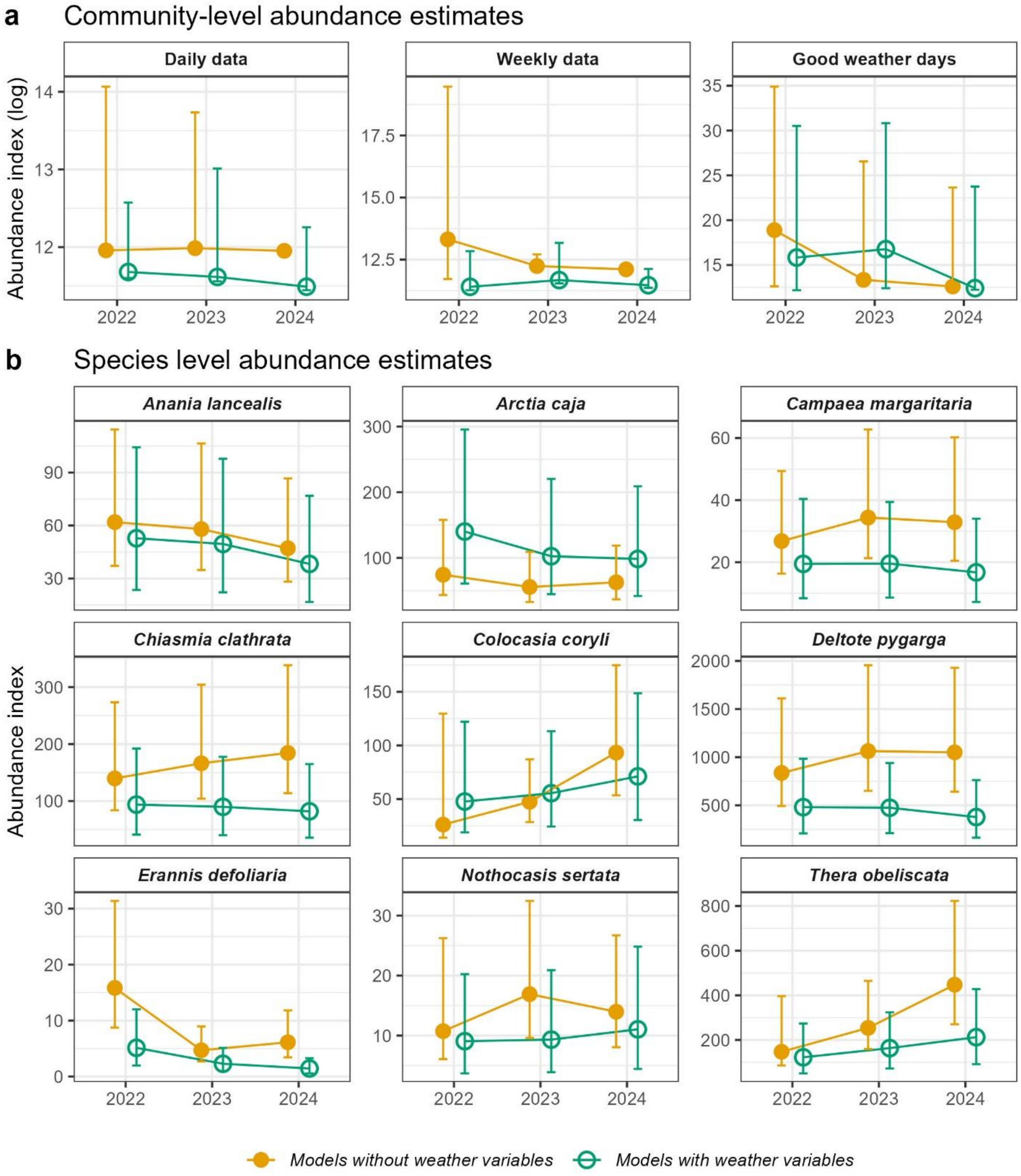


Figure 2: Abundance index estimates derived from GAMs for a) the community comprised of the

most abundant 100 species with identification precision above 80%, and b) selected species that show a range of sensitivity to weather conditions. For the community level graph (a), points are median abundance indices computed across 1000 posterior draws, and error bars are 95% credible intervals (see Box 1). The panels represent the full day-resolution dataset, and the Weekly and Good weather subsampled datasets. Note the different y-axes. Abundance indices were estimated using the area under the curve method for each species, trap, site and year before summing estimates to derive an overall community abundance during the trapping period. Species selected in b) all have identification accuracy above 86%.

The nine patterns in Fig. 2b represent the range of differences in abundance estimates between models with and without weather variables and suggest that it can be an important species- and research-specific decision in monitoring data analysis. Weather affects moth activity and detectability in the short-term influencing trap catches independent of population size, and through longer-term effects on survival and reproduction resulting in real abundance fluctuations [42,61]. Models with weather variables can account for daily or even hourly variation in activity, producing smoother seasonal patterns with lower uncertainty, but may underestimate abundance during years with favourable conditions that support larger populations. Conversely, models excluding weather capture realised population dynamics that are partly produced by weather-driven fluctuations but do not allow the separation of activity patterns and true abundance changes and may overestimate abundance on favourable days. This exploration highlights that the treatment of weather variables in abundance modelling warrants careful consideration based on research objectives. In any case, relative abundance estimates from fine resolution datasets could be invaluable additions to the monitoring toolbox if camera trapping is employed over long time periods, providing opportunities to compensate for weather fluctuations in moth monitoring and to link long-term abundance to fine-scale weather patterns to advance insights into moth ecology.

3.3 Species richness

Many biodiversity monitoring programmes also focus on providing information on diversity metrics, such as species richness, which can indicate how entire communities respond to environmental change [62] and is likely to contribute positively to community stability and ecosystem function [44]. For example, moth species richness has been shown to increase under less intensive agricultural practices (reviewed by [16]), and species richness losses in Hungary have been linked to land use and climate changes [62]. In Finland, species richness increases have been reported alongside general declines in abundance, probably reflecting a climate driven poleward expansion of geographical ranges [47,48]. These examples highlight the need to monitor both diversity and abundance across a wide network of sites.

To ensure comparability of species richness records across sampling units, variations in sampling effort must be accounted for using methods such as rarefaction and extrapolation [63], which estimate species richness at equivalent levels of sampling effort (Box 2). While high-resolution automated monitoring data minimises the need to rely on these methods, technical issues affecting any of the cameras can lead to variations in sampling effort. We therefore used rarefaction here and focus on the data of species that were verified by a Danish moth taxonomist (449 species). However, rather than only comparing species richness estimates across three biodiverse sites in Denmark, we also used this method to explore the estimates between different sampling frequencies with the aim of identifying an ideal balance between recording effort and estimation accuracy. To do this, we repeated the rarefaction process for different sampling regimes including our Weekly sample from above, as well as a dataset generated by subsampling every three days. We found that increasing intervals between sampling days led to greater levels of uncertainty in species richness estimates in all years, and underestimates of the number of species (Fig. 3a). The Weekly data estimate even re-ordered the sites in terms of the most diverse.

The reasons for the variability in uncertainty may be partly explained by the estimation method and its sensitivity to the number of species recorded during only one or two trapping days (Box 2). However, an illustration of the amount of information collected under the differing sampling regimes throughout the season also highlights the benefits of regular samples (Fig 3b). Under daily sampling, considerably more information about species occurrence is collected compared to less regular samples, and when a GAM is fit to species number with a smoother for day of the year, the uncertainty around the smoothed line is much reduced. Although the heights of the lines do not differ dramatically in Fig. 3b, species identity is not captured here and many species may be missed or underrepresented with longer sampling intervals, particularly at the peak of activity. The richness distribution curves in Fig. 3b also reiterate the point that continuous sampling through the season is likely to provide more complete estimates than snapshots (single sampling occasions in a season), and that a missed datapoint due to technical faults during weekly sampling is likely to be more costly than during daily sampling. Notably, however, sampling every third day provided relatively comparable estimates of diversity, which may be of interest for AMI trap users concerned with the impact of artificial UV light, the attraction of predators of nocturnal insects such as bats, birds, spiders and wasps and the overall storage and processing time of daily data.

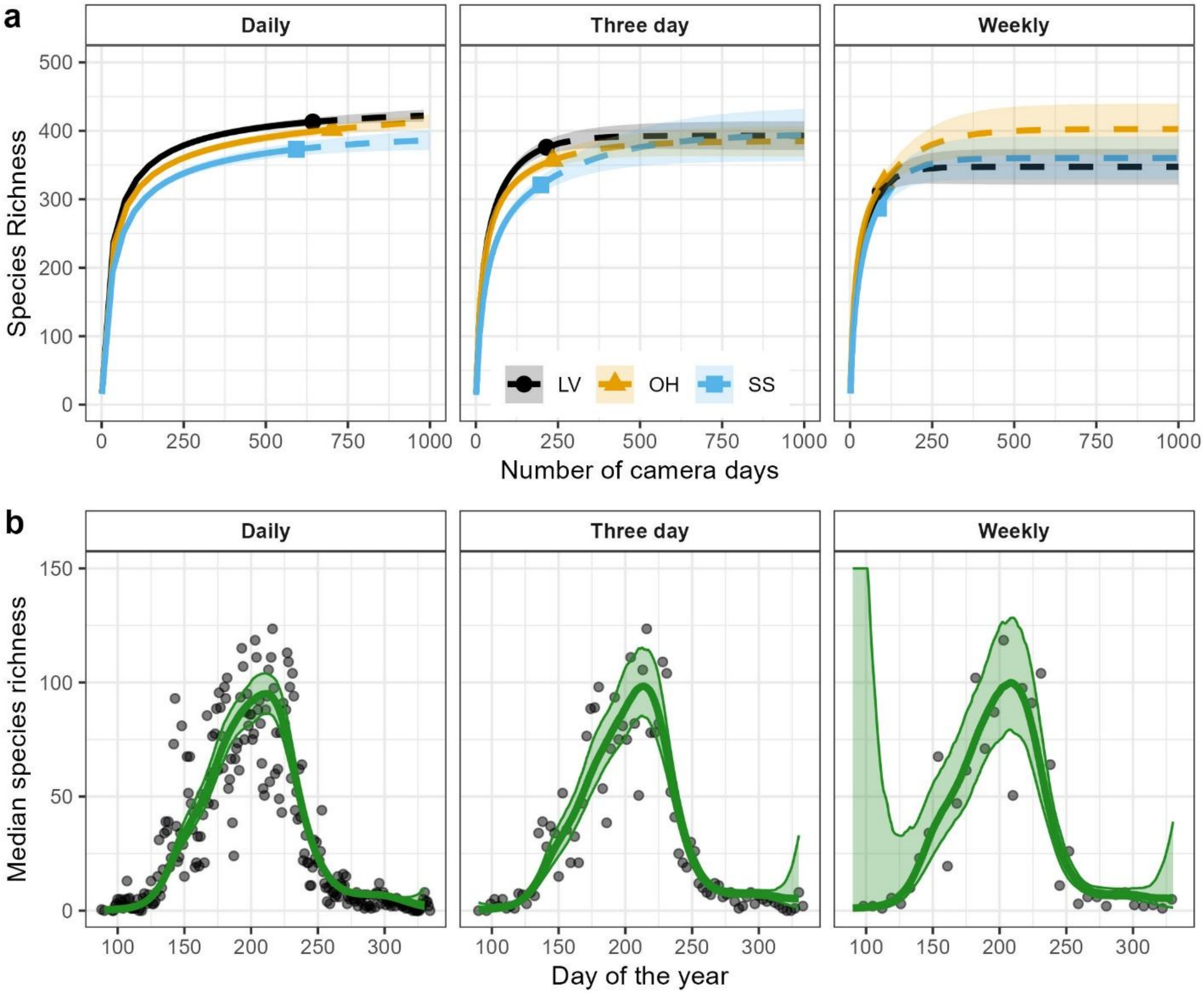


Figure 3: Moth species richness analysis results: a) species accumulation curves for the three sites in 2024 only, under different sampling regimes (Daily = data from every day, Three day = data from every third day, Weekly = data subsampled weekly). The point on each curve represents the actual sampled species richness. To compare sites at equal sampling effort (camera days), requires interpolation (solid line) or extrapolation (dashed line); b) the number of species

captured by the three different sampling regimes. Points are median species richness values across all sites and years, curves are GAM smoothers of richness over time with 95% credible intervals based on 1000 posterior samples. The relatively large error ribbon for the early part of the season of the weekly sampling occurs due to a lack of data.

3.4 Community composition dynamics

While abundance and diversity are popular metrics to communicate and interpret the state of an ecosystem or habitat, they provide only a partial view of the condition of an ecological community. In the context of climate change, ecologists are particularly concerned about the prospect of biohomogenisation: the process by which ecological communities become more similar across regions due to the dominance of common or generalist species [62,64]. For example, in Mediterranean forest reserves, moth community structure is linked to local weather conditions and vegetation structure, with dietary specialist species being especially sensitive to vegetation characteristics [43]. Similarly, long-term community analysis in Hungary has shown that species with narrow ecological niches are more likely to be replaced by larger and more generalist species [62]. These findings underscore the importance of monitoring changes in community structure and the underlying drivers of species co-occurrence, beyond basic diversity metrics. Furthermore, with international restoration targets such as the Nature Restoration Regulation requiring EU member states to restore 30% of degraded habitats by 2030 (European Commission 2024), advanced monitoring tools will be required to track the progress of these large scale ecological changes [65,66].

In light of these considerations, we explored the potential of automated sampling to generate informative data on community assemblages. For simplicity, we focused on data from two sites: Lille Vildmose and Ovstrup Hede, collected during 2024, and selected only those species that were present for at least 5 days and met arbitrary classification precision criteria (see Methods for details). This resulted in a pool of 142 species. Rather than employing the conventional approach of pooling data by site to compare overall community composition, we opted to analyse temporal patterns in community development (Box 3). To simplify visualization and analysis, we pooled data into weeks and then analysed the trajectory of community change throughout the season for the two sites; that is, the rate at which the relative abundance of species for each week changes through the season [67]. Our results demonstrated that the two communities diverged and converged at different points of the season, but with LV lagging behind OH during early summer before "catching up" in late summer (Fig. 4a). This temporal nuance would be obscured by single-time-point sampling or whole-season aggregation. Furthermore, the trajectories of two community similarity curves were almost circular, with early spring (April) and late autumn (October) communities showing greater resemblance to each other than to mid-summer communities (i.e., July & August = Weeks 28 to 35 = Day of the year 181 to 243). Early spring and late summer communities are likely to resemble each other due to the presence of "multivoltine" species (those with more than one flight period in a year) and species that are persistently present throughout the year (Fig 4b & c). Conversely, those groups showing strong seasonal peaks drive the turnover in species composition during the richest part of the year. This cyclical pattern was consistent for all three years (not shown), and longer term data could be further examined to explore changes in the timing of community "states" between years and their drivers [67]. These findings highlight the value of regular, automated sampling in capturing dynamic community processes, and emphasize that temporal turnover in community composition is an inherent feature of moth communities in this region and possibly more generally. This needs to be taken into consideration when drawing conclusions about spatial or interannual comparisons or when comparing the composition of communities before and after management interventions.

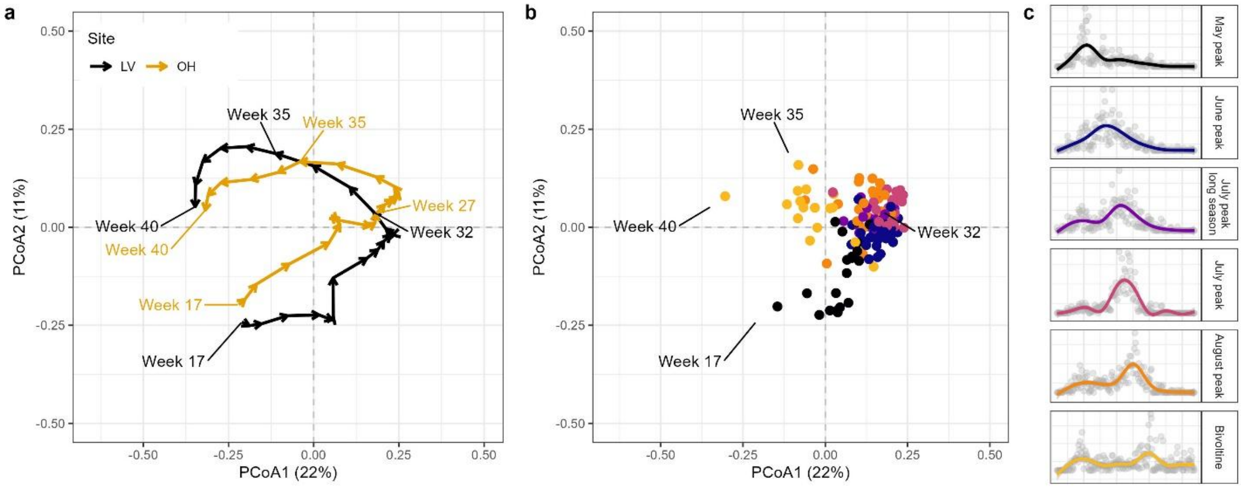


Figure 4: a) Community trajectory analysis of the 2024 moth communities of LV (Lille Vildmose) and OH (Ovstrup Hede), summed by week and across traps (a), and weighted average species scores plotted in the same ordination space (b). Species are coloured by their membership of a seasonal cluster (c). In ecological trajectory analysis, each week's community is treated as a separate ecosystem "state" in a Principal Coordinates Analysis (PCoA), and the changes from one week to the next are analysed geometrically. Thus, the length of an arrow in a) represents the speed of change in community states (i.e., the rate of species turnover). Arrowheads close together in ordination space are similar in composition, and those far apart are dissimilar. Although PCoA does not automatically compute species scores, the weighted average

scores can be interpreted in a similar way: species close together in b) are likely to be associated with same community states, and with the states in the same area of plot a).

4. Summary and future outlook

We demonstrate that automated insect monitoring systems can transform ecological monitoring and deepen our understanding of community and species-level dynamics. With the high temporal resolution of the data, and the potential for large spatial coverage, these systems can deliver data that provide more precise estimates of phenological events, more accurate estimates of abundance, standardised measures of diversity and detailed insights into community dynamics. These metrics are of fundamental importance to standardised monitoring schemes [5], and to track the success of restoration efforts [65]. With appropriate upscaling, automated systems could meet the criteria for cost-effective generation of indices fulfilling criteria for EBVs [25]. With data collection covering the entire season across a range of weather conditions, automated systems provide monitoring schemes with statistical power to overcome shortfalls of traditional systems, such as accounting for weather variability and avoiding the misleading focus on single time-point sampling. When combined with relevant abiotic information such as weather and land use variables, automated monitoring can also enable the modelling of future scenarios under potential ecological restoration and/or climate change with unprecedented precision.

Beyond statistical power, automated systems can confer other advantages. The AMI trap presented here and elsewhere [40,41] offers the opportunity to collect less destructive data on a wide range of species, and to deliver monitoring information while preserving human resources. Furthermore, the system can provide a standardised method of data collection with the potential to fit into a scientific sampling design, with automated data entry and quality assured data [14], and possibilities for spatial, inter- and intra-annual, and within-sample replication [5]. The system can be further enhanced to provide timely reporting of species and community trends over time using appropriate statistical approaches [14].

The approach presented here does not resolve all analytical challenges associated with insect camera trap data such as the multiple counting of individuals and the potential for over-representation of more active species compared to more sedentary species. More research is therefore required to quantify these potential biases and to understand the links between species activity, observed abundance and the role of weather conditions. Generalisation to other regions also relies on the availability of suitable training data, but the resulting datasets provide a remarkable level of detail that can only be matched by committed and extensive long-term monitoring programmes. This level of resolution opens up a wide range of monitoring and research possibilities that can assist with understanding trends, assessing management interventions, addressing fundamental and applied research questions and contributing to mechanistic modelling for future projections [2]. These possibilities can benefit new monitoring schemes, but also complement existing systems [14].

Insect populations and communities are extremely variable and dynamic, and recent evidence of alarming declines should be tempered by the knowledge that responses are heterogeneous [1,68].

This diversity of responses is important to capture across large spatial and temporal scales to better understand biodiversity changes and provide suitable conservation solutions [5,14]. Automated systems such as the AMI trap form a solution with an exciting level of detail and potential, and, as complimentary features in the monitoring toolbox, they will dramatically enhance biodiversity monitoring and our ability to meet conservation priorities in the future [24,27].

Methods

Phenology and abundance

To model the community and species level phenology patterns, we used a Generalised Additive Model (GAM) with a negative binomial error distribution, fit in the R programming environment [v4.3.0, 69] with the package *mgcv* [70] and thin plate regression spline as the smoothing basis. As these models are computationally intensive, we selected 100 species from the full dataset of >1600 species that met the following criteria: 1) species classification had been verified by a Danish moth taxonomist; 2) a total of more than 10 images were used in testing the precision of the identification algorithm, 3) precision was more than 80%, 4) species were present in all years for at least 5 days of camera trap recording. From the remaining species, we selected the 100 most abundant.

The phenology GAM used abundance as the response variable, and as explanatory variables we used a factor smooth for day of the year (doy) with year as the factor, to model the “global” community phenology curve, and smoothers for doy for each species to model species-specific phenology [71]. In full models including weather variables, we also fit smoothers to wind speed, relative humidity and temperature. These variables were recorded at each site by a weather station featuring the ZL6 advanced data logger (Meter, Munich, Germany).

We also included a “residuals covariate” to capture the daily relative activity of all moth species following the procedure described in Lindén et al. [60]. Briefly, to calculate the residuals covariate, a separate GAM was fitted for each species to the daily abundance of all other moth species (i.e. excluding the focal species), with a smoother for doy, factor smooth interactions (bs = “sz”) for doy and year and doy and trap location, and a random effect for site. From this model, log-scale residuals were calculated as

$$x_t = \ln[M_t + 1] - \ln[E(M_t) + 1]$$

where M is the abundance of all other moths on the *t*th day, and E(M) is the fitted value for M from the GAM. As described by Lindén et al. [60], this covariate can be seen as an alternative to a lagged autocovariate to account for autocorrelation, and describes relative moth activity in that those days with above average activity will have positive values, and days with below average activity have negative values. The inclusion of this covariate improved overall model fit of the phenology GAM and is likely to capture variation not explained by the three chosen weather variables above.

Finally, in the phenology GAM we included random effects for species and trap location. The model was fit with the *bam* function of the *mgcv* library with the fREML method, and k was set to 20 to ensure we captured the complexity of all seasonal patterns.

To make predictions with this phenology GAM, we computed a prediction grid that represented realistic weather conditions throughout the year, rather than use a constant for each variable (e.g. mean temperature, wind speed, humidity) that may provide unrealistic predictions during cooler or warmer periods of the year. To do this, we first fit GAMs to each weather variable (including the residuals covariate). Each weather model had random effects for trap location and year. We then made predictions of each variable for every day of the study periods, across all combinations of species, traps, sites and years. These predictions were combined into an overall prediction grid consisting of 686,600 observations and variables representing all those in the phenology GAM.

For predicting the community level phenology curves, we used this grid to draw 1000 samples from the posterior distribution of the model using a Gaussian approximation [72], using the *fitted.samples* function from the *gratia* package [73], providing 1000 abundance estimates for each species, trap, sites, year and day of the year. We aggregated these estimates across species, traps and sites to provide a 1000 total abundance values per day per year. For each day in each year, we then extracted the median total abundance to draw the predicted phenology curve, and the 0.025 and 0.975 quantiles to draw the 95% credible intervals. The community peak phenology date for each year was calculated both as the median date when peak abundance occurred in the samples and the date when 50% of the cumulative abundance predictions were reached.

For species level phenological metrics we used the same posterior samples, but for each species we calculated the day of the year when 10% (onset), 50% (peak) and 90% (end) of the cumulative abundance predictions were reached. In addition, we used the trapezoid rule to estimate relative abundance as the area under the curve using the *trapz* function from the *pracma* package [74]. These predictions for peak, onset, end and abundance were then summarised for each year by taking the median value of all fitted samples as the point estimate, and the 0.025 and 0.975 quantiles as the 95% credible intervals.

The modelling and predictions were repeated for a reduced phenology GAM, to compare the predictions for a model that does not include weather variables. The model structure was the same as above, except that it did not contain temperature, wind speed, relative humidity or the residuals covariate (which may represent other weather conditions important to moth activity). Finally, the two model types and the prediction process were repeated for two subsampled datasets for comparison of predictions between the full daily recorded observations, and observation sets based on reduced sampling capacity. The first of these was a “Weekly” dataset, which comprised observations subsampled from the full set every seven days, starting with the first day of sampling. The second was a “Good weather” dataset, which comprised all observations that met the following criteria during the four hours of sampling: temperature above 12 ^{o}C, wind speed below 1.5 m/s and precipitation below 1 mm. These values and variables were chosen to represent the kind of weather conditions that observers may choose to sample in to maximise data collection.

Species richness analysis

To analyse patterns in species richness we used the data for those species verified by a Danish moth taxonomist (449 species). To compare estimates of species richness across differing sampling regimes, we created two subsampled datasets: a weekly dataset as described above, and a "Third day" dataset, comprising observations subsampled from the full set every three days, starting with the first day of sampling.

For each of these samples, the number of sampling units was calculated as the number of days of working camera data collection for each site and year. Therefore, camera days were summed across the four traps at each site. The sampling units and species x site and year matrix were then inputs to the *iNEXT* function of the *iNEXT* package [75], which computes species diversity metrics (including species richness) across a range sampling intensities, utilising interpolation at intensities lower than the actual sampling effort, and extrapolation at higher intensities. The data type was "incidence frequency", confidence intervals (CIs) were set to 0.95 and endpoint was set to 1000 for each data set to provide comparable plots.

To plot species richness patterns over time, we first calculated species richness per day for each site and year (i.e. we combined data to count the number of species observed at each site), and again used the three comparison datasets – daily observations, every third day and weekly observations. These data were used as input for three separate GAM models, with richness as response variable, and a cubic regression spline smoother fit to Day of the year, allowing the smoother to interact with year, and for any site level deviations using the "sz" basis. The models were fit with the negative binomial error distribution. For each model, we then fitted the response (richness) to 1000 samples drawn from the posterior distribution (Gaussian approximation) using a full prediction grid. This provided 1000 estimates of richness for each day, site and year. We then averaged the richness estimate across site and year to derive 1000 estimates of species richness across the entire dataset for each day. The median richness for each day was then plotted as the best estimate of species richness, and 0.025 and 0.975 quantiles were plotted as 95% credible intervals.

Community composition

To initially group species by seasonal activity profile for visualisation purposes, we used hierarchical clustering of distribution patterns using the *hclust* function of the base R package [v4.3.0, 69]. To select species for this process, we used the same criteria as above (Expert verified, 10 images, 80 % precision, present for > 5 days), and for simplification we used only those species occurring at the LV and OH sites in 2024 (142 species). To estimate seasonal activity, we built a GAM for each species, with Day of the year (DOY) fitted as a thin plate regression spline, and trap location as a random effect, and the negative binomial error distribution. We then made abundance predictions for each day, location and species, and averaged abundance across traps to provide a mean abundance per day per species. These predictions were combined into a species x DOY matrix, which was then normalised with the *decostand* function of the *vegan* package [76].

This normalised matrix was converted to a Euclidean distance matrix with the *vegdist* function of the *vegan* package. The Ward.D hierarchical clustering method was used to identify seasonal activity clusters for each species. We experimented with the number of clusters by visually plotting species from the same cluster and found that 6 clusters provided the best balance between simplicity and adequately representing meaningful seasonal clusters.

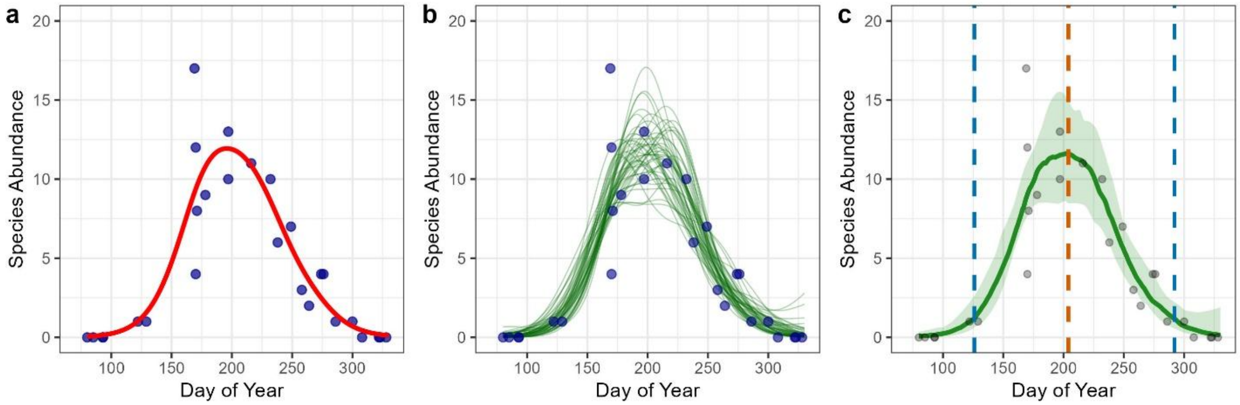
The same 142 species were used in ecological trajectory analysis (ETA) using the *ecotraj* package [67]. For simplicity, and to provide the longest period for comparison, we again focussed on the data from 2024 and the sites Lille Vildmose and Ovstrup Hede, which had at least two traps running continuously from week 17 to week 40. We first summed the data by week and site to aid visualisation, and then prepared a species x site/sampling week abundance matrix. For the ETA we therefore had two sites and 24 survey samples. The species x site matrix was normalised and a distance matrix was computed using the Canberra method. We then defined the trajectories using this distance matrix, the sites and survey times, and ran a trajectory Principle Coordinates analysis (PCoA). The first PCoA axis explained 22% of the variation in community "states" and the second explained 11%. Finally, to aid visualisation, we smoothed the trajectories using the *smoothTrajectories* function of the *ecotraj* package, which uses a Gaussian kernel to perform multivariate smoothing [67]. Finally, to visualise species within the PCoA ordination, we calculated Weighted Average Scores using the *wascores* function of the *vegan* package, as PCoA does not compute species scores.

Box Text

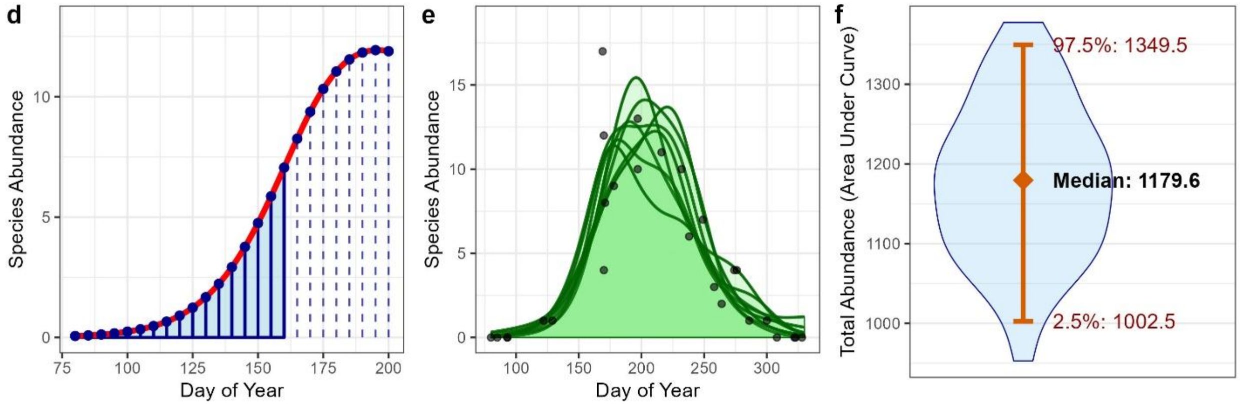
Box 1 – Daily data improve phenology and abundance estimates

Generalised Additive Models (GAMs) enable modelling of non-linear patterns between variables in a flexible way. This is particularly useful when modelling phenological patterns, where a species' seasonal activity may not conform to a simple parametric shape. The process used to estimate phenological events is depicted in the graphs below.

In panel a), the randomly generated abundance of a single hypothetical species is plotted against day of the year (DOY), and a single smoothed line has been fit that best captures the non-linear nature of the pattern. Panel b) shows the same abundance points but displays the curves of 20 "bootstrap" samples – each green line is based on a random resample of the data with replacement. Replicated many times (typically 500 or 1000), the sampled smoothed lines provide a distribution of predictions, which converge on the most likely scenario given the data and the specified model. Each sample can provide a separate estimate on the date of peak activity, and other phenological events such as the likely onset of activity (when 10% of all observations have been made) and end of activity (90 % of all observations). Panel c), shows how the confidence band (shaded in green) represents the range of bootstrapped samples, while the median of the peak dates across bootstrap samples is selected as the most likely (the red dashed line). Similarly, the blue dashed lines represent the most likely onset and end dates. "Credible intervals" of these dates are represented using the 2.5% and 97.5& quantiles across bootstrap samples. If more data is available in graph a), the bootstrapped samples in b) are more likely to resemble each other and uncertainty will be low. Thus, the high-frequency data provided by daily AMI trap operation will provide much more information for predictions than less frequent sampling.



The abundance curves can be used to generate robust estimates of abundance in a given year. In panel d), a section of the curve from a) is in focus and each day has been separated into a polygon or “trapezoid”. The area of each day’s trapezoid can be estimated using integration, and the areas for all days summed to provide an abundance estimate. Again, uncertainty can be addressed by bootstrapping the curve, and repeating the abundance calculation for each sample (e). The most likely abundance value, shown in f), is then the median of all the bootstrapped sample estimates, while the 2.5% and 97.5% quantiles provide the credible interval range.

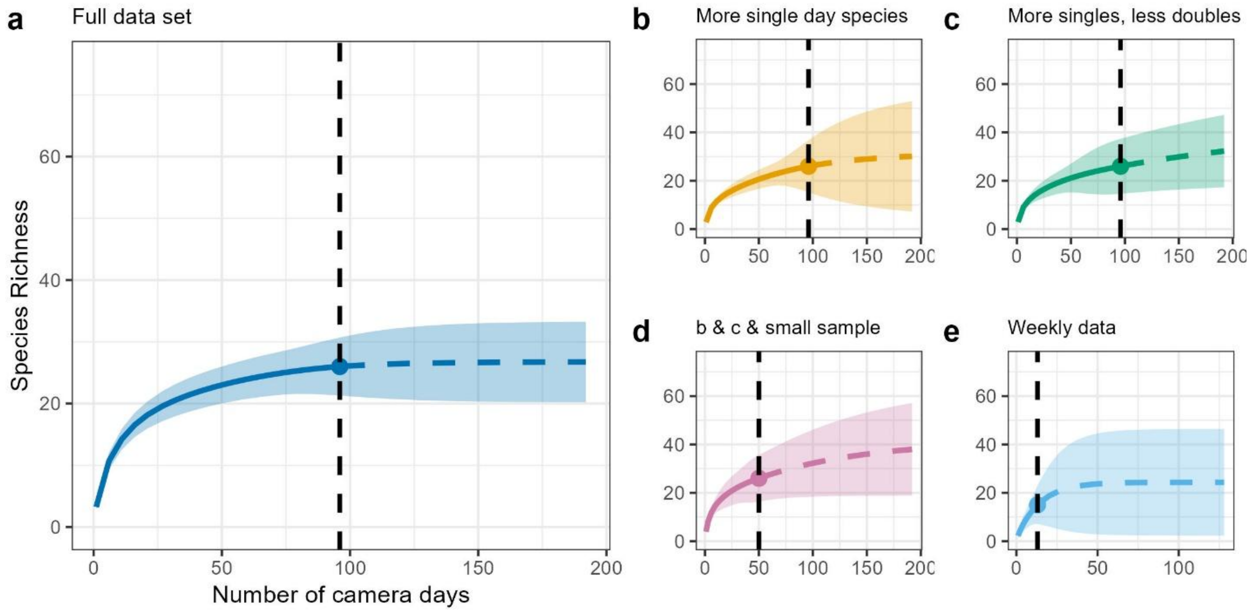


As with other regression techniques, the addition of covariates will alter the location and smoothness of the prediction lines and allow predictions to be made for differing levels of the covariates. Hence, the addition of e.g., weather variables in the GAM of species phenology may alter the size and shape of the fitted curve and therefore the relative abundance calculation.

Box 2 – Daily monitoring data can improve rarefaction curves

Estimating and comparing species richness from multiple ecological communities can be challenging, particularly under varying sampling effort (e.g. the number of traps, survey days, survey area, individuals encountered). Most species in a community are likely to be rare and many will go undetected, and the number of species encountered will increase non-linearly with sampling effort up to an unknown asymptotic value. To solve this problem and estimate unbiased values for species richness in multiple assemblages at equivalent levels of sampling effort, Colwell et al. [77] and Chao and Jost [78] unified a framework of statistically modelling species richness using so-called rarefaction curves. The process enables the comparison of species richness at a common reference point. If the reference point is a smaller number of samples than was collected, interpolation or "rarefaction" is used to reduce the estimate down to a random subsample. If the reference point is greater, extrapolation is used following the trajectory of the rarefaction curve.

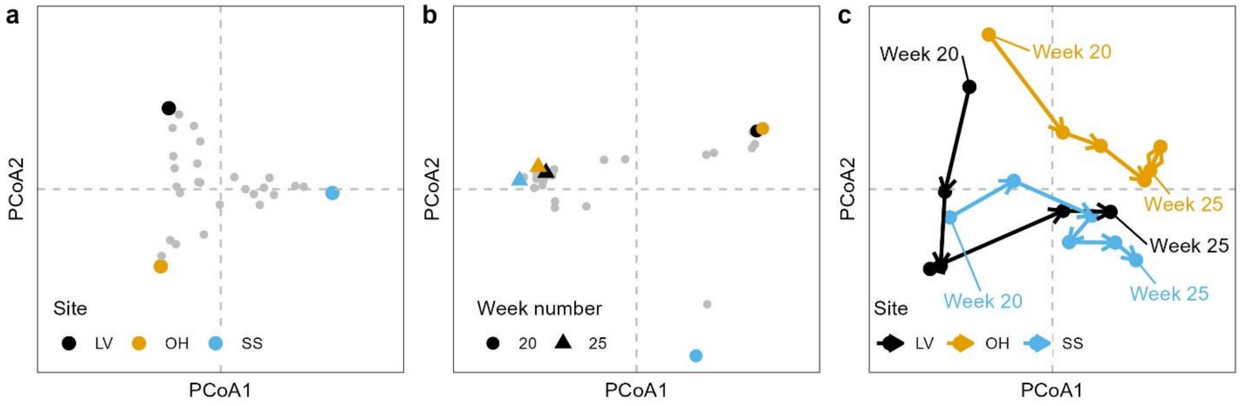
However, uncertainty estimates are susceptible to sampling issues and we demonstrate this here with a small subsample of the AMI trap data. Panel a) shows the species richness curve for 30 randomly selected species from one site during a 30-day period in 2024. This period provided 96 camera days across the four traps, and 28 out of 30 species were detected during this time (the dashed vertical line). The species accumulation curve provides estimates of the number of species detected with fewer or more camera days. The uncertainty around the curve highlights that some rare species may have been missed during sampling (the upper boundary of the shaded area) while some may have been over-represented (the lower boundary). These uncertainty estimates are sensitive to the number of "singletons" in the data (species only captured once); in panel b) the same data are used but 5 species have been altered: instead of occurring on 3 days they only occur on 1 day (More single day species). The number of "doubletons" (species found twice) and the ratio between singletons and doubletons is also an important factor: in graph c), "More singles, less doubles", the same dataset from b) is used but four doubletons have been changed to occurring on 3 days. Thus, fewer doubletons and a higher ratio increase uncertainty values. A third issue is shown in d), "b & c & small sample", where the same data from c) are combined with a low sample size (only 50 camera days). Uncertainty is driven up here as small samples of very diverse communities increase the "guesswork" involved in estimating unseen diversity. Finally, when the sample of data from a) is sub-sampled weekly instead of daily (resulting in 16 camera days – 4 weeks x 4 traps), the resulting graph is shown in e), "Weekly data". Here, the large uncertainty area is due to a combination of all three of the above issues. Therefore, daily sampled data provide more information about the true abundance distributions of both common and rare species. Note also that these findings are for a given daily sampling window (11pm to 3am) with motion detection enabled to detect movement at 2 second intervals. With lower sampling rates, detection probability would also be reduced.



Box 3 – Daily data enable Community Trajectory Analysis

In many studies of community composition, the objective is to understand how similar or dissimilar species assemblages are and why. Communities made up of the same species in comparable abundances are considered “similar” and will be plotted close together in two dimensional “ordination plots”. Communities made up of different species are dissimilar and will be far apart in ordination space, and coplotting values of likely predictors such as climate, pollution or habitat can help to determine the factors behind dissimilarity. In a basic sense, community composition data are often pooled across sampling periods and plotted together to give an overview of a assemblages observed across a season. However, below we demonstrate with a small subsample that this approach may mask important community dynamics. The data featured here comprise all species from the traps at all three study sites but limited to weeks 20 to 25 in 2024. Panel a) shows a traditional ordination, using data pooled across this 6-week period, showing the similarities in the make-up of their moth communities. The grey dots are the species, and those close to each other and to particular sites are also associated. Here it appears that each site is relatively distinct. However, given that we have sampled the sites many times over the season, we could analyse the communities at different times separately. In panel b), we compare the sites at two distinct points in time, pooling only week 20 and week 25. Here, all three sites are relatively similar in week 25, but site SS is distinct from the other two in week 20. This may suggest some hidden dynamics between these time periods, or some unknown event leading to a dramatic shift in composition.

Community trajectory analysis aims to unmask these patterns, treating repeated samples of the same sites as different community “states” and then assessing the speed and direction of change over time. In panel c), this analysis has been performed for the three sites, pooling the data for each week to provide 6 community states each. Here, longer arrows indicate large changes in community composition and short arrows indicate the opposite. The direction of arrows also denotes divergence or convergence of community similarities. For example, site LV starts and ends with a similar community to site OH, but diverges in weeks 22 and 23. In later analysis, we might delve deeper into the data to understand why, but such investigation can only be conducted with data collected frequently and throughout the season. Continuous monitoring of these sites will also reveal whether such patterns recur, become more extreme or disappear over time. Similarly, further analyses can help define how appropriate comparisons of community composition across space and time should account for within-season variability.



References

1. Wagner, D. L. Insect Declines in the Anthropocene. *Annual Review of Entomology* **65**, 457–480 (2020).
2. Besson, M. *et al.* Towards the fully automated monitoring of ecological communities. *Ecology Letters* **25**, 2753–2775 (2022).
3. Bohmann, K. *et al.* Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution* **29**, 358–367 (2014).
4. Lindenmayer, D. B. & Likens, G. E. Adaptive monitoring: a new paradigm for long-term research and monitoring. *Trends in Ecology & Evolution* **24**, 482–486 (2009).
5. Montgomery, G. A., Belitz, M. W., Guralnick, R. P. & Tingley, M. W. Standards and Best Practices for Monitoring and Benchmarking Insects. *Front. Ecol. Evol.* **8**, (2021).
6. Butchart, S. H. M. *et al.* Global Biodiversity: Indicators of Recent Declines. *Science* **328**, 1164–1168 (2010).
7. Dirzo, R. *et al.* Defaunation in the Anthropocene. *Science* **345**, 401–406 (2014).
8. Hallmann, C. A. *et al.* More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLOS ONE* **12**, e0185809 (2017).
9. Burkle, L. A., Marlin, J. C. & Knight, T. M. Plant-Pollinator Interactions over 120 Years: Loss of Species, Co-Occurrence, and Function. *Science* **339**, 1611–1615 (2013).
10. Losey, J. E. & Vaughan, M. The economic value of ecological services provided by insects. *Bioscience* **56**, 311–323 (2006).

11. Potts, S. G. *et al.* Safeguarding pollinators and their values to human well-being. *Nature* **540**, 220–229 (2016).
12. CBD. Decision adopted by the conference of the parties to the convention on biological diversity 15/4. Kunming-montreal global biodiversity framework. (2022).
13. Kremen, C., Ullman, K. S. & Thorp, R. W. Evaluating the Quality of Citizen-Scientist Data on Pollinator Communities. *Conservation Biology* **25**, 607–617 (2011).
14. Pocock, M. J. O. *et al.* Developing and enhancing biodiversity monitoring programmes: a collaborative assessment of priorities. *Journal of Applied Ecology* **52**, 686–695 (2015).
15. Dunn, R. R. Modern Insect Extinctions, the Neglected Majority. *Conservation Biology* **19**, 1030–1036 (2005).
16. Fox, R. The decline of moths in Great Britain: a review of possible causes. *Insect Conservation and Diversity* **6**, 5–19 (2013).
17. Farley, S. S., Dawson, A., Goring, S. J. & Williams, J. W. Situating Ecology as a Big-Data Science: Current Advances, Challenges, and Solutions. *BioScience* **68**, 563–576 (2018).
18. Kissling, W. D. *et al.* Towards a modern and efficient European biodiversity observation network fit for multiple policies. https://ecoevorxiv.org/repository/view/7993/ (2024).
19. Van Dooren, T. J. M. Assessing species richness trends: Declines of bees and bumblebees in the Netherlands since 1945. *Ecol Evol* **9**, 13056–13068 (2019).
20. Warren, M. S. *et al.* The decline of butterflies in Europe: Problems, significance, and possible solutions. *Proceedings of the National Academy of Sciences* **118**, e2002551117 (2021).

21. Groenendijk, D. & van der Meulen, J. Conservation of moths in The Netherlands: population trends, distribution patterns and monitoring techniques of day-flying moths. *Journal of Insect Conservation* **8**, 109–118 (2004).
22. van Klink, R. *et al.* Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science* **368**, 417–420 (2020).
23. Høye, T. T., Montagna, M., Oteman, B. & Roy, D. B. Emerging technologies for pollinator monitoring. *Current Opinion in Insect Science* **69**, 101367 (2025).
24. van Klink, R. *et al.* Emerging technologies revolutionise insect ecology and monitoring. *Trends in Ecology & Evolution* **37**, 872–885 (2022).
25. Jetz, W. *et al.* Essential biodiversity variables for mapping and monitoring species populations. *Nat Ecol Evol* **3**, 539–551 (2019).
26. Pereira, H. M. *et al.* Essential Biodiversity Variables. *Science* **339**, 277–278 (2013).
27. Høye, T. T. *et al.* Deep learning and computer vision will transform entomology. *Proceedings of the National Academy of Sciences* **118**, e2002545117 (2021).
28. Pimm, S. L. *et al.* Emerging Technologies to Conserve Biodiversity. *Trends in Ecology & Evolution* **30**, 685–696 (2015).
29. Marquina, D., Esparza-Salas, R., Roslin, T. & Ronquist, F. Establishing arthropod community composition using metabarcoding: Surprising inconsistencies between soil samples and preservative ethanol and homogenate from Malaise trap catches. *Molecular Ecology Resources* **19**, 1516–1530 (2019).
30. Roslin, T. *et al.* A molecular-based identification resource for the arthropods of Finland. *Molecular Ecology Resources* **22**, 803–822 (2022).

31. Doi, H. *et al.* Detection of an endangered aquatic heteropteran using environmental DNA in a wetland ecosystem. *Royal Society Open Science* **4**, 170568 (2017).
32. Noguerales, V. *et al.* Community metabarcoding reveals the relative role of environmental filtering and spatial processes in metacommunity dynamics of soil microarthropods across a mosaic of montane forests. *Molecular Ecology* **32**, 6110–6128 (2023).
33. Kohlberg, A. B., Myers, C. R. & Figueroa, L. L. From buzzes to bytes: A systematic review of automated bioacoustics models used to detect, classify and monitor insects. *Journal of Applied Ecology* **61**, 1199–1211 (2024).
34. Roy, D. B. *et al.* Towards a standardized framework for AI-assisted, image-based monitoring of nocturnal insects. *Philosophical Transactions of the Royal Society B: Biological Sciences* **379**, 20230108 (2024).
35. Mann, H. M. R. *et al.* Automatic flower detection and phenology monitoring using time-lapse cameras and deep learning. *Remote Sensing in Ecology and Conservation* **8**, 765–777 (2022).
36. Collett, R. A. & Fisher, D. O. Time-lapse camera trapping as an alternative to pitfall trapping for estimating activity of leaf litter arthropods. *Ecology and Evolution* **7**, 7527–7533 (2017).
37. Möglich, J. M. *et al.* Towards reliable estimates of abundance trends using automated non-lethal moth traps. *Insect Conservation and Diversity* **16**, 539–549 (2023).
38. Keitt, T. H. & Abelson, E. S. Ecology in the age of automation. *Science* **373**, 858–859 (2021).

39. Bjerge, K., Karstoft, H. & Høye, T. T. Towards edge processing of images from insect camera traps. *Remote Sensing in Ecology and Conservation* **n/a**, (2025).
40. Bjerge, K., Nielsen, J. B., Sepstrup, M. V., Helsing-Nielsen, F. & Høye, T. T. An Automated Light Trap to Monitor Moths (Lepidoptera) Using Computer Vision-Based Tracking and Deep Learning. *Sensors* **21**, 343 (2021).
41. Jain, A. *et al.* Insect Identification in the Wild: The AMI Dataset. in *Computer Vision – ECCV 2024* 55–73 (Springer, Cham, 2025). doi:10.1007/978-3-031-72913-3_4.
42. Neff, F. *et al.* Moth communities are shaped by season, weather, elevation, and landscape composition. *Insect Conservation and Diversity* **18**, 670–680 (2025).
43. Uhl, B., Wölfling, M. & Fiedler, K. Exploring the power of moth samples to reveal community patterns along shallow ecological gradients. *Ecological Entomology* **47**, 371–381 (2022).
44. Seymour, M. *et al.* Ecological community dynamics: 20 years of moth sampling reveals the importance of generalists for community stability. *Basic and Applied Ecology* **49**, 34–44 (2020).
45. Conrad, K. F., Warren, M. S., Fox, R., Parsons, M. S. & Woiwod, I. P. Rapid declines of common, widespread British moths provide evidence of an insect biodiversity crisis. *Biological Conservation* **132**, 279–291 (2006).
46. Groenendijk, D. & Ellis, W. N. The state of the Dutch larger moth fauna. *J Insect Conserv* **15**, 95–101 (2011).

47. Antão, L. H., Pöyry, J., Leinonen, R. & Roslin, T. Contrasting latitudinal patterns in diversity and stability in a high-latitude species-rich moth community. *Global Ecology and Biogeography* **29**, 896–907 (2020).

48. Ellis, E. *et al.* Extinction in the north and colonisation in the south: the latitudinal drivers of community warming. Preprint at https://doi.org/10.21203/rs.3.rs-4899409/v1 (2024).

49. Isaac, N. J. B., van Strien, A. J., August, T. A., de Zeeuw, M. P. & Roy, D. B. Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* **5**, 1052–1060 (2014).

50. Rolnick, D., Bundsen, M., Jain, A. & Cunha, F. RolnickLab/ami-data-companion. RolnickLab (2023).

51. Thackeray, S. J. *et al.* Phenological sensitivity to climate across taxa and trophic levels. *Nature* **535**, 241-U94 (2016).

52. Visser, M. E. & Gienapp, P. Evolutionary and demographic consequences of phenological mismatches. *Nat Ecol Evol* **3**, 879–885 (2019).

53. Altermatt, F. Tell me what you eat and I'll tell you when you fly: diet can predict phenological changes in response to climate change. *Ecology Letters* **13**, 1475–1484 (2010).

54. Maurer, J. A. *et al.* Phenological responses of 215 moth species to interannual climate variation in the Pacific Northwest from 1895 through 2013. *PLOS ONE* **13**, e0202850 (2018).

55. Végvári, Z. *et al.* Life-history traits and climatic responsiveness in noctuid moths. *Oikos* **124**, 235–242 (2015).

56. Wu, S. & Shiao, M. Temporal and forest-type related dynamics of moth assemblages in a montane cloud forest in subtropical Taiwan. *Journal of Asia-Pacific Entomology* **26**, 102073 (2023).
57. Rafferty, N. E., CaraDonna, P. J., Burkle, L. A., Iler, A. M. & Bronstein, J. L. Phenological overlap of interacting species in a changing climate: an assessment of available approaches. *Ecol Evol* **3**, 3183–3193 (2013).
58. Schmucki, R. *et al.* A regionally informed abundance index for supporting integrative analyses across butterfly monitoring schemes. *Journal of Applied Ecology* **53**, 501–510 (2016).
59. Didham, R. K. *et al.* Interpreting insect declines: seven challenges and a way forward. *Insect Conservation and Diversity* **13**, 103–114 (2020).
60. Lindén, A., Meller, K. & Knape, J. An empirical comparison of models for the phenology of bird migration. *Journal of Avian Biology* **48**, 255–265 (2017).
61. Jonason, D., Franzén, M. & Ranius, T. Surveying Moths Using Light Traps: Effects of Weather and Time of Year. *PLOS ONE* **9**, e92453 (2014).
62. Valtonen, A. *et al.* Long-term species loss and homogenization of moth communities in Central Europe. *J. Anim. Ecol.* **86**, 730–738 (2017).
63. Chao, A. *et al.* Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs* **84**, 45–67 (2014).

64. Carvalheiro, L. G. *et al*. Species richness declines and biotic homogenisation have slowed down for NW-European pollinators and plants. *Ecology Letters* **16**, 870–878 (2013).
65. Menéndez-Miguélez, M. *et al*. How to measure outcomes in forest restoration? A European review of success and failure indicators. *Front. For. Glob. Change* **7**, (2024).
66. Vörös, M. *et al*. Towards harmonised monitoring of grassland restoration: a review of ecological indicators used in the temperate region. *Journal of Arid Environments* **230**, 105426 (2025).
67. De Cáceres, M. *et al*. Trajectory analysis in community ecology. *Ecological Monographs* **89**, e01350 (2019).
68. Høye, T. T. *et al*. Nonlinear trends in abundance and diversity and complex responses to climate change in Arctic arthropods. *Proceedings of the National Academy of Sciences* **118**, e2002557117 (2021).
69. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2023).
70. Wood, S. N. *Generalized Additive Models: An Introduction with R*. (CRC Press., 2017).
71. Pedersen, E. J., Miller, D. L., Simpson, G. L. & Ross, N. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ* **7**, e6876 (2019).
72. Wood, S. N. Simplified integrated nested Laplace approximation. *Biometrika* **107**, 223–230 (2020).
73. Simpson, G. L. _gratia: Graceful ggplot-Based Graphics and Other Functions for GAMs Fitted using mgcv_. (2024).

74. Borchers, H. _pracma: Practical Numerical Math Functions_. (2023).
75. Hsieh, T. C., Ma, K. H. & Chao, A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution* **7**, 1451–1456 (2016).
76. Oksanen, J. *et al.* vegan: Community Ecology Package. (2019).
77. Colwell, R. K. *et al.* Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* **5**, 3–21 (2012).
78. Chao, A. & Jost, L. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* **93**, 2533–2547 (2012).

Acknowledgements

This research was funded by Aage V. Jensen Naturfond grant number N2024-0013 and the European Union's Horizon Europe Research and Innovation programme, under Grant Agreement No. 101060639 (MAMBO).